ELSEVIER

Artificial Intelligence 72 (1995) 53–79

Artificial Intelligence

Sensorimotor transformations in the worlds of frogs and robots*

Michael A. Arbib, Jim-Shih Liaw*

Center for Neural Engineering, University of Southern California, Los Angeles, CA 90089-2520, USA

Received September 1992; revised July 1993

Abstract

The paper develops a multilevel approach to the design and analysis of systems with "action-oriented perception", situating various robot and animal "designs" in an evolutionary perspective. We present a set of biological design principles within a broader perspective that shows their relevance for robot design. We introduce schemas to provide a coarse-grain analysis of "cooperative computation" in the brains of animals and the "brains" of robots, starting with an analysis of approach, avoidance, detour behavior, and path planning in frogs. An explicit account of neural mechanism of avoidance behavior in the frog illustrates how schemas may be implemented in neural networks. The focus of the rest of the article is on the relation of instinctive to reflective behavior. We generalize an analysis of the interaction of perceptual schemas in the VISIONS system for computer vision to a view of the interaction of perceptual and motor schemas in distributed planning which, we argue, has great promise for integrating mechanisms for action and perception in both animal and robot. We conclude with general observations on the lessons on relating structure and function which can be carried from biology to technology.

1. An evolutionary background

Agents do not have an unlimited repertoire of behavior. An animal evolves within a certain ecological niche, and even that "general-purpose" animal, the human, builds on its evolutionary heritage to acquire specialized skills that fit the person for a relatively limited set of roles in society. Similarly, when we write a

* The research described in this paper was supported in part by grant no. 1RO1 NS 24926 from the National Institutes of Health and Grant NOOO14-92-J-4026 from ONR (M.A. Arbib principal investigator).

* Corresponding author. E-mail: liaw@rana.usc.edu.

SSDI 0004-3702(94)00055-7

computer program or build a robot, we design it to carry out a certain repertoire of tasks. The animal or the robot comes with specialized receptors which allow it to sense only a subset of the energies available in the environment, and has specialized effectors, including those which can mediate active sensing. (For a formal characterization of evolutionary specialization based on environmental constraints see Horswill [20].)

We will focus upon a particular animal, the frog (and its close cousin, the toad) to construct a biological "robot" which will allow us to see more carefully the way in which a specific biological system has evolved—but our task will be to understand the integrated style of action and perception that such evolution yields, rather than to chart the evolutionary process itself. We start with an analysis of aspects of the basic survival behaviors of feeding and fleeing—approach and avoidance—and then model the control structures of the animal either as a network of interacting automata-like systems, which we call *schemas* (or in some cases, schema instances) [4], or as neural networks which implement the functionality of these schemas.

The work in the next three sections is part of a general research program called *Rana computatrix* (the frog that computes), an evolving testbed for multilevel modeling (in terms of both schemas and neural networks) of the mechanisms of visuomotor coordination [2]. Frog and toad are sufficiently similar that our generic model applies to both of them. Thus, except when referring to specific experiments carried out with one or the other animal, we will use the term "frog" throughout when the generic frog/toad is meant. While the core constraint of the *Rana computatrix* study is to understand the biological data, the work is at the same time designed to yield case studies which can point us to a more general understanding of principles of action-oriented perception of the animal/robot acting in the world.

2. Schemas for approach and avoidance

To simplify rather drastically, we may say that the frog's ability to find food and escape enemies can be reduced to the ability to tell small moving objects from large ones. A frog surrounded by dead flies will starve to death, but the frog will snap with equal "enthusiasm" at a moving fly or a pencil tip wiggled in a fly-like way. On the other hand, a larger moving object can trigger an escape reaction. Thus, at a very simple level, we might imagine that the brain of the toad has two basic pattern-recognition routines (what we shall call *perceptual schemas*), one for recognizing small moving objects (food-like stimuli) and one for recognizing large moving objects (enemy-like stimuli). We could then come up with the very simple model shown in Fig. 1(a) in which we have the signals from the eye routed to these two perceptual schemas. If the small-moving-object schema is activated, it will in turn trigger the *motor schema* (our term for an automaton or control system for controlling action) to get the animal to approach what is apparently its prey; while if the perceptual schema for large moving objects is activated, it will

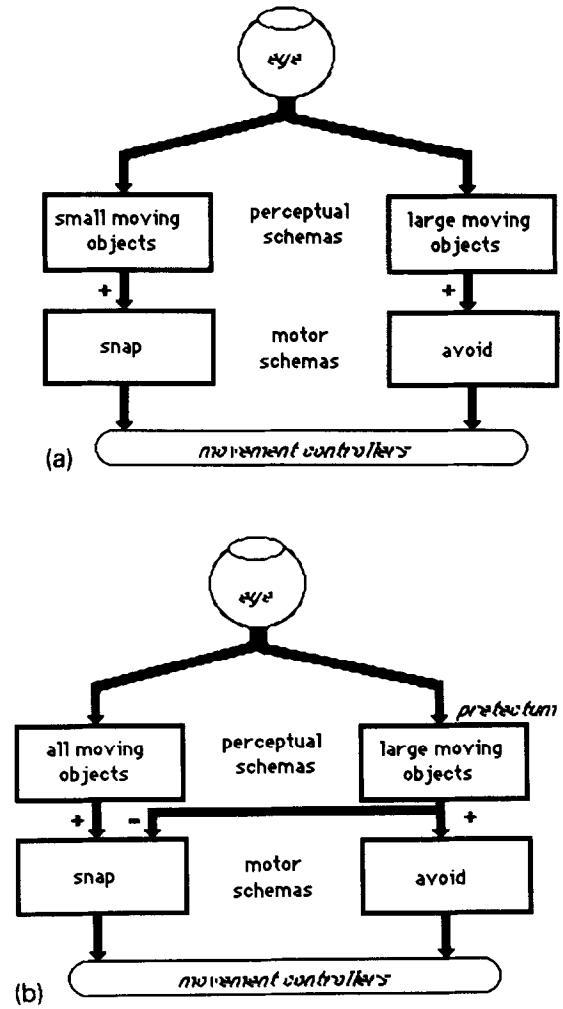

Fig. 1. (a) The "naive" schema program for the toad's snapping and avoidance behavior. (b) The schema program revised in light of data on the effect of lesioning the pretectum.

trigger the motor schema for avoidance, causing the animal to escape an apparent enemy.

Before going further, we may note that the system in Fig. 1(a) involves four simple automata, and Lyons and Arbib [34] have modeled schema instances as port automata in applying them to perceptual robotics. However, rather than simply process input symbols to yield output symbols, the individual automata have *activation levels* which measure some degree of confidence, and it is the more active of the two perceptual schemas that will trigger the appropriate motor schema to yield the appropriate response. We may say that the perceptual schemas *compete* to control the behaviour of the animal. This is a very simple

example of the type of mechanisms of competition and cooperation that can be exhibited by a network of schemas. In particular, multiple motor schemas may be coactivated to control subtle behaviors [11].

Notice the important point that perceptual schemas do not serve simply as YES–NO recognizers, equipped with a confidence level, but rather serve to provide a parametric description which can be used in tuning motor behavior appropriately. When it recognizes prey, the animal does not respond with a launch in a standard or random direction, but rather it snaps at the position in 3-D space where the prey is located. Similarly, as we shall see in more detail in Section 4, when the animal sees a predator-like stimulus, it must find an escape direction that will have a good likelihood of taking it out of the path of the predator.

We now consider how such a model can become a biological model as distinct from a purely functional model. Our assertion will be that a schema-based model (we turn to neural network models in a later section) becomes a biological model when explicit hypotheses are offered as to how the constituent schemas are played over particular regions of the brain. For the examples to be offered in this paper, some simple anatomy will suffice: The right eye of the frog projects to the left half of the brain, with the most important visual midbrain target being the *tectum*. The projection from the retina to the layered structure of the tectum preserves the neighborhood relationships of the retina (and thus the visual field), and is called a *retinotopic* map. We will also study the *pretectum* (so called because it is in front of the tectum). Returning to the schemas of Fig. 1(a), we may note (without giving the details) that experimentalists had reason to suggest that the small-moving-object schema was in the tectum, while the large-moving-object schema was in the pretectum. Such hypotheses about neural localization may be tested by lesion experiments. The model of Fig. 1(a) predicts that animals with a pretectal lesion would continue to approach small moving objects just as the normal animal would, but would not respond at all to large moving objects.

However, the model is false! Peter Ewert studied toads (see [15] for a review) in which the pretectum had been lesioned. He found that not only did the toads respond to small moving objects with approach behavior, but they also responded to large moving objects with approach behavior! This observation leads to the new schema-level model shown in Fig. 1(b). The new data tell us that in the absence of the pretectum, the animal must be able to respond to all moving objects with approach. Thus, we replace the left-hand perceptual schema for *small* moving objects by a perceptual schema for *all* moving objects. On the other hand, in the normal animal, we have that recognition of large moving objects triggers avoidance and so we leave the right-hand column the way it was. However, although we have now explained the response of the lesioned animal to all moving objects, and the response of the normal animal to large moving objects, it remains to tune the model so that the normal animal will respond to small moving objects with approach but not avoidance. This we can achieve by having an inhibitory pathway running from the perceptual schema for large moving objects (in the pretectum) to the approach schema—or, equivalently, to

the schema for all moving objects. With this model we can now explain our small database on the behavior of both normal animals and those with a lesion of the pretectum. We have thus established that hypotheses about neural localization of subschemas may be tested by lesion experiments. As we shall see below, we may then proceed further to model a brain region to see if its neural circuitry will indeed implement the posited schemas. The important point is that biological models can be expressed at the level of a network of interacting "automata", and that these can really be biological models that can be tested at the level of such a coarse-grained network, whether or not data or hypotheses are available about the fine-grain implementation of those automata in neural networks.

With this background, we are well on our way to understanding our first principle:

Principle 1 (*Cooperative computation of schemas*). The functions of perceptual-motor behavior and intelligent action of animals and robots situated in the world can be expressed as a network of interacting schemas/schema instances. The method of interaction of schemas/schema instances is "cooperative computation" (Competition/Cooperation) so that "computations" which are often seen as the province of traditional symbol-based processing are carried out by distributed "neuron-like" methods which do not involve explicit symbolic control. This not only serves as a basis for coordinated motor actions, but for reactive planning, and intelligent behavior, including the use of language.

In biology, we must discover how the schemas are distributed across biologically distinct brain regions. In robotics, the schema analysis and the allocation of processors may occur in tandem as part of an explicit design process (see Section 8).

3. Detours and path planning

In the present section, we show how a study of frog behavior led to a model of path planning that was then applied to the design of the control of mobile robots, and which anticipated an important area of robot control. The motivation for this work is provided by Ingle's study [22] of a frog observing a worm through a semi-transparent grating or barrier. Instead of launching directly at its prey as would occur if no barrier is present, the frog reacts appropriately, detouring around the barrier to get its prey. We note here how this behavior leads us to extend the schema picture of Fig. 1(b). Now, the perceptual schema for recognizing prey must be augmented by the perceptual schema for recognizing a barrier, and there can no longer be a simple direct path from prey recognition to the triggering of approach behavior. Rather, there must be some way for this path to be modulated by the recognition of the barrier to yield an indirect detour, rather than the direct response.

In the particular situation under consideration here, the animal must not only

recognize prey and barrier, but must locate them in space. If it can recognize that the prey is in front of the barrier or at most a tongue's length behind, then the animal will indeed snap directly. But if the prey is further behind the barrier, then the animal must use its recognition of where the prey is and where the barrier is to come up with a path which will carry it around the barrier towards the prey.

Arbib and House [6] offered a model of this in which perceptual schemas for prey and barrier can drive motor schemas which compete and cooperate to yield the overall behavior (Fig. 2). We postulate that the ground plane in front of the animal is represented in the brain, with the worm being represented as a global attractor: each point in the "arena" has a vector pointing towards the worm, with the vectors decreasing in length with distance from the prey, but not vanishing (Fig. 2(a)). On the other hand, each fence post is represented as a local repellor, such that animal will be repelled either to left or right if it comes close to the fence post, but will not be affected if it is further away (Fig. 2(B)). When we combine the action of the individual fence posts, we get a strong vector field to the left of the post and a strong vector field to the right. If we combine all this activity we get the field shown in Fig. 2(c) which can be integrated to yield trajectories which either pass to the left or to the right of the fence and then continue en route to the prey.

It is not our claim here that that the brainstem of the frog implements the above potential field algorithm in its neural circuitry. Rather, the crucial point is that we have an evolutionary account of how such a system might arise (Fig. 3): The elements of the prey-recognition system—perceptual schema, motor schema, and motor pattern generator (MPG)—co-evolve so that activity in the prey-schema can represent a goal in such a way that the approach-schema provides the right control signals for the MPG to determine a path to the prey; while the detour system evolves (or co-evolves) by combining a perceptual schema for stationary objects with a motor schema that *modulates* the effect that the approach-schema has on the motor pattern generator.

In building upon this approach, Arkin [7] developed a hybrid architecture for the control of a mobile robot. The robot was equipped with a map of its "world" and, given a knowledge of its current position and of where it was to go, it could plan a path.[1] What Arkin did was to translate such a path into a vector field which consists of vectors pointing along the path plus "diagonal" vectors in the neighborhood of the path pointing the robot back onto the central path in a direction tending towards that of the path. He then added to the system a "frog brain" to navigate around obstacles—not actually grafting a biological brain onto the control computer of the robot, but augmenting the AI planner with a

[1] In fact, this planner was more "symbolic" than "neural". The study of animals rests on the expectation that all schemas will be implemented in neural networks, or will be expressed in the functional interactions of the neural networks and the biomechanical systems they control—i.e., they will represent the function of a dynamic system not all of which need be neural. However, in an artificial system, different implementations may be optimal for different schemas, and for many tasks symbolic processing will marry well with available VLSI technology.

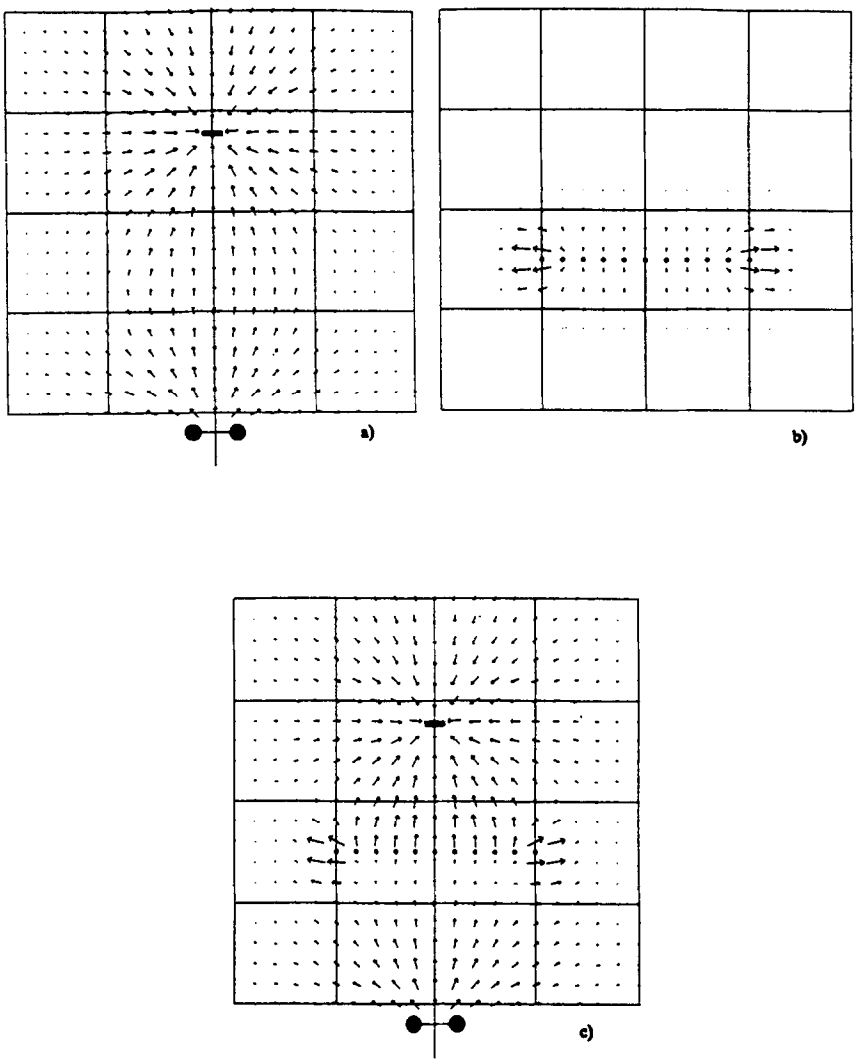

Fig. 2. The potential fields depicted here represent an exploratory attempt at defining a set of primitive fields which will interact in interpreting a complex scene. Each field provides, for each position in the ground plane, a vector showing the direction and "strength" of a movement the animal might make were it situated at that position. (a) A single prey object sets up a radially symmetric attractant field whose strength decays gradually with distance from the prey. (b) A single barrier object sets up a repellant field whose effect is more localized to its point of origin than is that of the prey field. The barrier field is not radially symmetric but has a lateral component that is stronger but decays more rapidly with distance than does its opposing component. (c) The effect of the interaction of the fields from several barrier objects arranged to form a fence is to provide a strong lateral thrust at the fence ends. The lateral components produced by the interior posts is effectively cancelled by neighboring posts. The net field produced by the interaction of all of the elements of the configuration can then be thought of as tracing out a set of paths, most of which are diverted around the fence ends. (Arbib and House [6]).

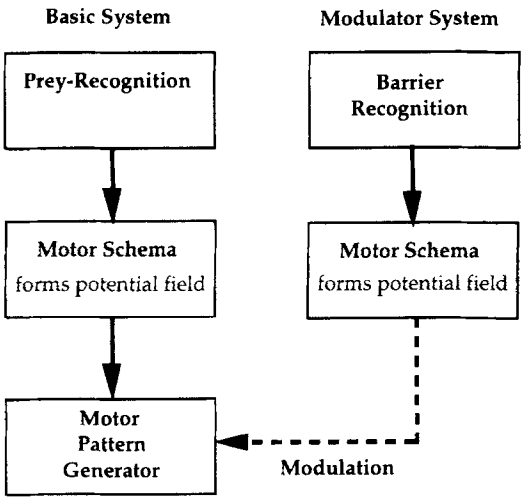

Fig. 3. The evolutionary refinement of adding detour behaviour to prey-approach by modulating the basic system for approaching prey.

potential field system similar to that in the Arbib–House study of the toad. When an obstacle was detected along the path, it set up a local repulsive field. The robot's path was then determined by the local field established by the path vectors and the obstacles. In general, the robot would follow the path, detouring to one side on detecting an obstacle. After being displaced toward the side and leaving the field of the obstacle, it would move back towards the center of the path and then continue following the path specified by the AI planner. Such a method has now been discussed many times as the "potential field" method for path planning in robotics (see, e.g., [25, 26]). See [38] for an alternative approach to path planning.

To close this section, we emphasize not so much the particular implementation of the schemas, but rather the evolving subtlety of the schema interactions. We start with two basic systems (Fig. 1(a)) for response to small and large moving objects, respectively, and then find that if we are to match the biological strategy we must come to the more subtle interactions shown in Fig. 1(b), in which recognition of small moving objects is not in fact localized in any one region, but is rather a system property involving the *modulation* of the tectum by the pretectum. We then extend the complexity of the environment to which the animal responds—it no longer contains a single prey or a single predator to which the animal may respond with the most basic forms of the "survival behaviors" of feeding or fleeing, but now contains the more subtle structure of obstacles which can block the animal's path and around which the animal must now be equipped to detour. In evolutionary terms, this corresponds to expansion of the ecological niche in which the animal is well suited to survive—just as Arkin's basic design for a mobile robot was extended from path-following in an uncluttered environment to work in a world which contains obstacles. We now have a sense of the new

perceptual schemas that must be added (for barrier detection) as well as the type of modulation that must be involved. This leads us to our next principle:

Principle 2 (*Evolution and modulation*). New schemas often arise as "modulators" of existing schemas, rather than as new systems with independent functional roles.

Further examples of Principle 2 may be seen in the lesion-based analysis of schemas for approach and avoidance behavior [11]. Here it simply suffices to note that this strategy for the analysis for biological systems may be traced back to the work of Braitenberg [8] and Walter [47], and has a number of points in common with the approach to the design of robot controllers offered by Brooks [9].

4. Neural mechanism of avoidance behavior

We now complete our introduction to biological analysis of "robots" by briefly outlining a specific example of how the schemas which serve animal behavior can be traced down to the neural networks that implement them.[2] Fig. 4 shows a number of escape behaviors exhibited by the frog. In Figs. 4(a)–(c) [22], we see the direction of approach of a large moving object, and we see bars radiating from the head of the frog providing a histogram of the relative frequency with which the animal chooses a particular escape direction over a number of trials. The escape direction may be briefly characterized as a compromise between the forward direction of the animal and the direction immediately away from the looming stimulus. Barriers can modify avoidance behavior, just as they modify approach behavior. If we use the data from Fig. 4(b) to determine the preferred direction of the animal for a looming stimulus coming directly from the left of the frog and interpose a barrier to block that preferred direction, then, as we see in Fig. 4(d), the behavior of the animal changes and it no longer tends to jump in the previously preferred direction, but just to the left or just to the right of the barrier. In the rest of the section, we will not return to the study of barriers, but will instead look in more detail at how the looming stimulus is recognized, and how this recognition is transformed into action.

Before doing so, we look at more recent experimental data [23] which show that the behavior is actually more subtle (Fig. 4(e)). In Fig. 4(a)–(c), we saw that if a stimulus is looming directly at the frog, its escape direction will be a compromise between the direction away from the stimulus and the forward direction of the frog. However, Fig. 4(e) shows that if, instead of moving directly towards the animal, the stimulus is on a trajectory which will carry it in front of

[2] Among the many biologically based neural network models have been developed as part of *Rana computatrix* are models of retina (Teeters and Arbib [44]), tectal–pretectal interactions in prey recognition (Cervantes-Pérez, Lara and Arbib [10]), and the role of anterior thalamus and medial pallium in habituation (Wang and Arbib [48]), as well as the model presented here.

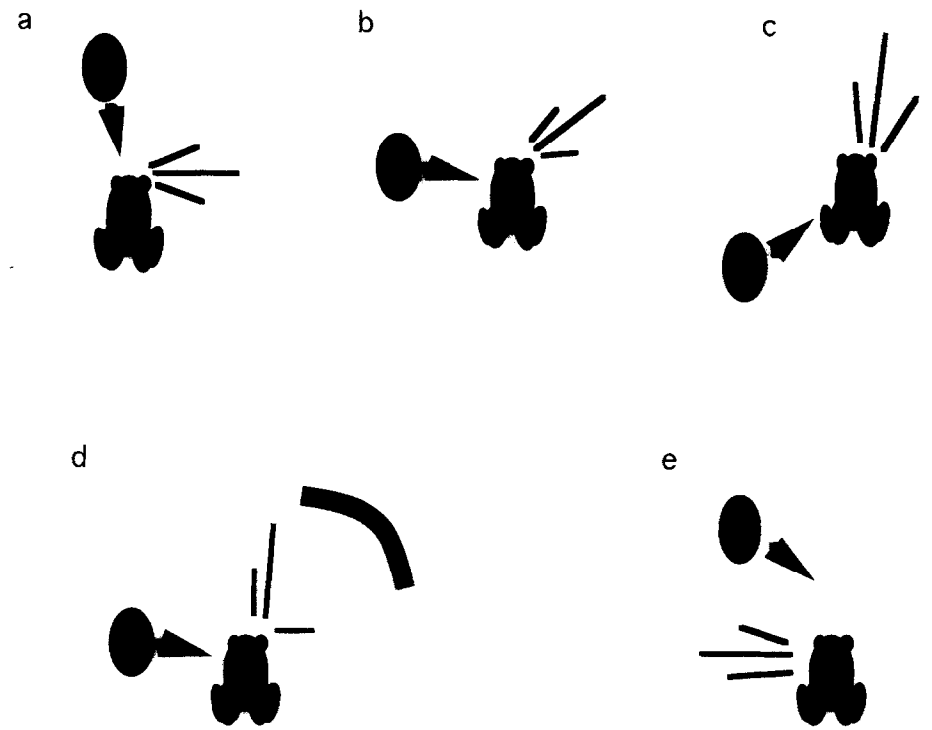

Fig. 4. The escape direction. When a looming stimulus is on a collision course with a frog, the escape direction of the frog is a compromise between the forward direction and that away from the looming object (a)–(c). However, if a barrier is interposed in the preferred direction, the animal will tend to jump to either side (d). If the stimulus is not on a colliding trajectory, the frog will jump in such a direction as to "cut back" behind the looming object (e). ((a)–(d) are adapted from Ingle [22] and (e) from Ingle and von Hoff [23]).

the animal, the frog will not respond based solely on the position of the looming stimulus, for in this case the animal would follow one of the trajectories shown in Fig. 4(a) which would carry it on a collision path with the predator, thus making it highly likely to be captured. Instead, the animal exhibits the highly adaptive "cutback" behavior shown in Fig. 4(e) which carries it on a path that is well away from that being pursued by the looming object. Here again we note an evolutionary refinement going from simple recognition of a predator to recognition of the trajectory that the predator is following to find an escape direction based on this extra information. Clearly, it is an interesting question to understand how evolutionary pressure could yield an animal able to exhibit this more subtle repertoire. However, in this section our concern is with the neural mechanisms which provide the frog with this functional repertoire. Unlike the previous sections in which we have simply sought a specification of schemas (whether or not analyzed in terms of specific brain regions) to provide a functional

analysis of the behavior, we now look at the availability of neurophysiological data.

To start the discussion going, we recall why the frog was initially seen as an interesting candidate for a "biological robot". In one of the classic papers on neural networks, Pitts and McCulloch [40] offered a hypothetical network for the recognition of "universals"—in other words, for recognizing an object despite great variations in the size, position or orientation with which it was projected on the retina. They postulated ways in which arrays of feature detectors could be gathered together in such a way that invariants with respect to a variety of group operations could be found which would then yield a pattern of neural activity which corresponded to the "universals" of a pattern, rather than its particular presentation. On this basis, some ten years later Oliver Selfridge (well known for his 1959 paper on Pandemonium [43]) persuaded Lettvin to look for such a structure in the frog's brain. The calculation of group invariants was not found to be exemplified, but arrays of feature detectors were indeed found. The ganglion cells of the retina, those which send their axons back to the brain, were classified in four different classes, and these four classes of cells were found to project to different depths in the tectum, forming four specific retinotopic maps. Even more excitingly, these maps appeared to be tied to the behavioral repertoire of the animal. If we call the four classes R1, R2, R3 and R4 (remember each one is a spatially arrayed population of cells in the "output layer" of the retina) we find that the R3 cells seem to respond best to small moving objects—as if they were bug detectors—whereas the R4 cells seemed to respond best to the large moving objects—as if they were enemy detectors [29]. The slightly later work of Hubel and Wiesel (e.g., [21]) on cat and then monkey found cells in visual cortex which reported on the orientation of edges, and thus provided very low-level features for shape description. By contrast, it appeared that Lettvin et al. had hit the jackpot in linking vision to action by showing that the frog retina computed features specifically linked to the behavioral repertoire of approach and avoidance, feeding and fleeing. For this reason, as well as the linkage to the pioneering modeling of Pitts and McCulloch, the frog was seen in the 1960s to be an excellent candidate for a neural network analysis of a visual system which was not simply engaged in some sort of abstract pattern processing but was directly geared to the determination of appropriate actions.

Fortunately or unfortunately, research over the subsequent years—both by neurophysiologists and by modelers—has shown that the story is not so simple, and that the frog really does need a brain (!) and not just a retina to determine whether to feed or flee. For example, it has been shown that the activity of the R3 cells is quite complicated. To a first approximation it can be viewed as response to the *leading edge* of a moving object entering its receptive field (the region of the visual field in which patterns of illumination can affect the activity of the cell), rather than the presence of a prey-like object within the receptive field. Again (and this was already recognized by Lettvin et al.) the activity of the R4 cell can be better interpreted as just a measure of *dimming*—but clearly a predator by

casting a larger shadow will dim the receptive field more than prey, and thus excite the R4 cell more strongly.

Many studies of visual processing or visually guided behavior ignore the particular transformations conducted by the retina, or simply reduce them to a contrast enhancement difference-of-Gaussians lateral inhibition mask. However, we [31] have taken as our starting point the properties of the R3 and R4 neurons. Recall that the R3 and R4 neurons (which have been modeled by Teeters and Arbib [44]) form an array which stretches across the output layer of the retina. If we consider the effect of a dark looming stimulus, we will then see greatest activity amongst those R3 cells whose receptive fields include the leading edge or expanding boundary of that looming stimulus, whereas the R4 cells which will respond most strongly are those contained within the interior of that expanding pattern. We can then combine these cells to provide a model of the T3 cells (so called because they are type 3 among cells of the Tectum as characterized by their physiological response). The T3 cell will respond more and more strongly as a stimulus looms to a position at the center of its receptor field if the connections from R3 cells to the T3 cell are radially symmetric but with a sort of inverted Gaussian form in which activity towards the periphery is more effective than activity at the center. The larger the looming stimulus, the further out the pattern of R3 activity and thus the stronger the input to the T3 cell. Since the T3 cells also form a retinotopic array, many cells will be activated by this looming stimulus but the T3 whose receptive field is centered on the center of the looming stimulus will have the strongest response. To complete the design of the T3 cell, and to make it responsive to a large looming stimulus, but not to several small objects flying apart from each other or to an expanding ring, we give the R4 cells a standard Gaussian projection to the T3 cells so that darkness at the center of the receptive field of the T3 cell will increase its response.

With this we have given analysis that does not simply explain how the presence of a looming stimulus can be represented by a peak of activity in an array of neurons, but also uses circuitry with cells whose firing rates provide a good model of firing rates actually observed neurophysiologically (see [31] for the details and simulations). We now have the neural network implementation of the perceptual schema required to explain the behavior shown in Figs. 4(a)–(c), where the current *position* of the looming stimulus on the retina implies the preferred direction of escape. However, to provide the necessary perceptual schema for the situation shown in Fig. 4(e), we must come up with cells that recognize the *temporal to nasal motion* across the retina. It turns out that the T2 neurons, also in tectum, do have this sensitivity. However, we have no neurophysiological data as to how these cells are actually wired up and so we use a standard model of directional selectivity to link an array of T3 neurons whose activity signals the current center of the looming stimulus feed to an array of T2 neurons in such a way that the passage of the stimulus from left to right will increase the likelihood of a T2 neuron firing. By contrast, if the pattern is moving in the opposite direction, then the direct activation of a T2 neuron by the corresponding T3 neuron will be diminished by the inhibition received via delay neurons to the

right. The resultant neural network will respond more vigorously to a pattern moving from left to right than to a pattern moving from right to left, within a given velocity range.[3]

With this we have all the perceptual information we need to complete our model. In Fig. 5(a) we see the retinotopic map of T3 neurons (simplified from a two-dimensional to a one-dimensional retina for ease of comprehension here) with an appropriate projection pathway from neurons in the T3 array to neurons in what we call the motor heading map which will cause the animal to turn towards the retinotopically corresponding location. Unlike the projection from prey-recognition neurons to the motor heading map, in which retinotopically corresponding points are linked by the projection, we now set up the connections to replicate the data of Figs. 4(a)–4(c) so that each peak of activity on the T3 layer will yield a peak of activity in the motor heading map centered at the point

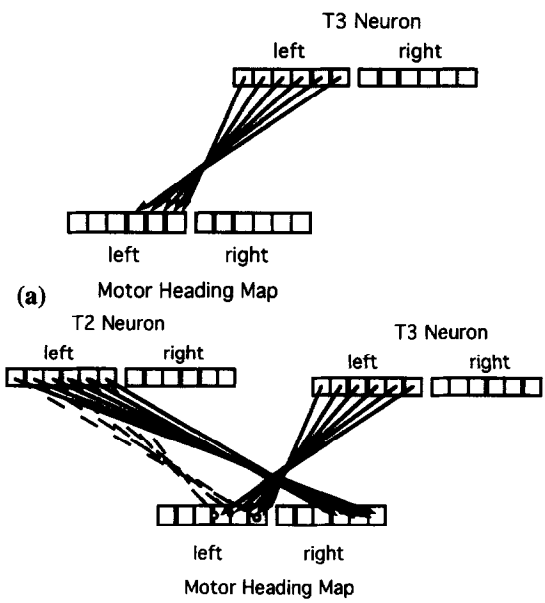

Fig. 5. Gating of the tectal projection onto motor heading map. Only half of the projections are shown here for simplicity. The T3 neurons project to the heading map to indicate the escape trajectory for looming object on a colliding trajectory. When the stimulus is crossing the visual field, the T2 signal blocks the T3 signals while exciting the contralateral heading map, thus resulting in a "cutback" jump.

[3] Direction-selective neurons in *mammalian* visual systems each have a preferred (small) range of velocity so that together they provide a sort of population coding of stimulus velocity. Motion-related neurons in the anuran (R2, R3, T2, T3, etc.), on the other hand, respond to the entire range of velocity. The stimulus velocity is coded in their firing rate—the greater the speed, the higher the firing rate. In terms of stimulus direction, most T2 cells prefer movement in the behaviorally significant temporal to nasal direction.

which is a compromise between the forward direction and the direction away from the looming stimulus. This will control the motor pattern generators to yield the Figs. 4(a)–(c) responses.

The situation must now be made more subtle to address the data of Fig. 4(e), and we show how this is done in Fig. 5(b). First, we add the T2 array which again is retinotopic but is now signaling not simply that there is a looming stimulus at that point of the retina but that this stimulus is moving from left to right. (We will not look at the corresponding connections for a stimulus on the right half of the visual field moving to the left—this will simply be the mirror image). This activity must block the normal escape response and so we see a projection from cells of the T2 array to the motor heading map which is a replica of the projection from the T3 array to the motor heading map—but this new projection is *inhibitory*. Thus, if there is no T2 activity the system will respond as in Fig. 5(a). To complete this extended model, we must now add a new excitatory pathway from the T2 neurons which project to the cutback direction, and thus can trigger the appropriate motor behavior in this case, while avoiding the normal escape behavior. What is perhaps worth stressing here is that the T3 neurons serve as the basis for the "evolution" of a more "sophisticated" set of neurons, the T2 neurons which not only signal the presence of a looming stimulus but the presence of a looming stimulus moving on a particular type of trajectory. This new system then *modulates* T3 activity by projecting to the same motor heading map and competing there with the T3 input, rather than acting upon the T3 array itself. This is another instance of Principle 2 on Evolution and Modulation.

We refer the reader to [31] for further details of the neurophysiological data which grounds the model, for further details of the circuitry involved in the model, and for a number of analyses of the simulations that we have conducted. Reflecting on the control schemes shown in Figs. 1, 3 and 5 we come to the third principle:

Principle 3 (*Interaction of partial representations*). A multiplicity of different representations—whether they be partial representations on a retinotopic basis, abstract representations of knowledge about types of object in the world, or more abstract "planning spaces"—must be linked into an integrated whole. Such linkage, however, may be mediated by distributed processes of competition and cooperation. There is no one place in the brain where an integrated representation of space plays the sole executive role in linking perception of the current environment to action.

We close the presentation by noting that simulation shows that as the stimulus speed increases, the time until the network yields a motor command decreases. This is a direct relation that follows from the network properties that we have sampled to explain the avoidance behavior. The velocity-dependent response time is due to the property of the T3 and other neurons whose firing rates increase as the speed of a looming stimulus increases. This result may be contrasted with the great body of work in the motion vision literature which interposes the computa-

tion of the optic flow field between the retinal input and the motor output. Where Gibson [17] noted that the optic flow could provide useful input for navigation, Lee and Lishman [26] explicitly showed how the time until contact of an approaching object could be inferred from the optic flow field, and suggested ways in which this explicit time parameter could be used to gate action. Here, we have offered an alternative model in which the action is controlled adaptively without the explicit extraction of time until contact from the optic flow.

5. Application of the looming avoidance model to robot control

The model for frog looming perception is directly applicable to obstacle avoidance in autonomous navigation, since, from a relative motion point of view, the situation where an object moves towards a robot is similar to one in which the robot moves towards the object. Moreover, the two situations become identical when a mobile robot has to negotiate moving obstacles. An experiment has been conducted to test the capability of the looming avoidance model in detecting obstacles and providing a detour path [30, 32]. The robotic experiment was carried out in an integrated testbed consisting of Neural Simulation Language (NSL) [49] for implementing the neural network model, and the Rapid Robotics Application Development environment (R^2AD) [16] for dynamic control of a robot arm. A camera mounted on the moving robot arm provides the visual inputs to the neural network, which computes the 3-D motion of the obstacles and determines an appropriate "escape" direction to guide the robot arm to go around the obstacles (Fig. 6).

Though the detection of obstacles and the computation of trajectory relative to the moving camera is performed in the same way as described above, there is one major difference between the looming avoidance behavior and obstacle avoidance. In the simulation of looming avoidance, only one approaching stimulus is presented at a time, whereas multiple obstacles are present in the robot experiment. This raises several interesting issues that the model must deal with, including occlusion, detection of gaps between obstacles, sizes constancy for obstacles of different sizes at different distances, and the interaction and integration of multiple obstacle signals.

Unlike avoiding a single looming object where the data of escape direction is obtained experimentally and the projection from the looming detectors is fixed, the heading for avoiding multiple obstacles has to be determined dynamically based on their spatial arrangement. Here we adopt the motor heading map [11] to provide a substrate on which signals of multiple obstacles interact and compete with each other and a heading of the next step for the robot to take emerges from such interaction. For this scheme to work, the signal should indicate not the location and extent of the obstacle, but rather, it should specify the opening beyond the edges of an obstacle. This is achieved by projecting looming detectors to the motor map via a connectivity patten that resembles an inverted DOG (Difference of Gaussians). Through such a convolution, neurons activated by an

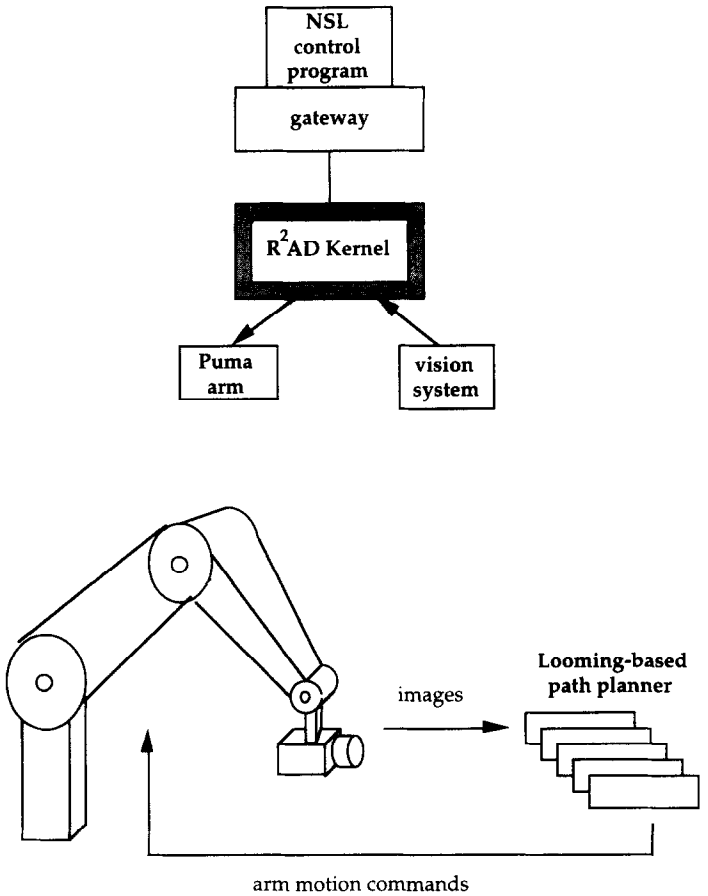

Fig. 6. Experimental setup for robot navigation. The illustration at the top shows the architecture of the integrated system. The high-level control is performed by neural network models written in NSL, whereas the low-level robot manipulation is computed by R^2AD. An interface (gateway) provides the protocol for communication between these two components. In addition to controlling the robot arm, R^2AD also provides frame grabbing routines for visual inputs from a camera. One image frame is taken after every move made by the robot and is fed into the looming perception neural network. The neural network locates the position of the obstacles based on the looming patterns and selects an optimal path based on some criteria (locally shortest path in this case) by specifying the direction of the next movement.

obstacle inhibit cells on the map that correspond to the same spatial location and excite those that are some distance away. As a result, only cells whose location corresponds to an opening (or gap between two obstacle) can be activated.

The presence of multiple obstacles (and hence multiple gaps) raises the question of which route to choose. A winner-takes-all mechanism proposed by Didday [13] is built into the motor map. A simple mechanism which gives higher

preference to more centrally located gaps is employed to obtain a locally optimal path (shortest path length). This is implemented as a differential tonic firing rate for cells in the motor map such that more centrally located ones have higher background activity. Therefore, a more centrally located gap has a greater probability to emerge as the winner on the map. Here we see that the motor map offers a medium to which various sources can send biasing signals (e.g., the desirability of a path or deviation from the intended route, etc.) to modulate the competition (cf. Principles 1, 2 and 3).

Since the only available information is the expansion of the image of obstacles, the model is confronted by a problem, namely, without depth information, how can the model respond to a small obstacle at a closer distance instead of a larger one at a greater distance when the latter subtends a bigger visual angle (i.e., it casts a bigger image on the camera)? Although only expanding patterns in the image are available, it provides a crucial clue about the distance of an object, namely, motion parallax. Motion parallax refers to the fact that for two objects moving at the same speed at different distances, the closer one generates a greater retinal shift, i.e., its image moves at a higher speed than the more distant one. Therefore, once the camera starts moving, the looming detector will respond more strongly to the smaller and closer obstacle since it elicits more retinal ganglion cells. In several experiments with obstacles of different sizes, the model demonstrated the capability of correctly avoiding them based on motion parallax (Fig. 7; for details of the robotic application see [32]).

6. From instinctive to reflective behavior

So far we have focused on instinctive behavior with sensorimotor transformation at its core. The sensorimotor system is composed of schema assemblages that are "evolutionarily hardwired" into patterns of competition and cooperation between specific brain regions. The animal performs in a "reactive" mode to interact with a dynamic environment, with little or no lookahead planning. Although great efficiency can be achieved via such hardwired schema assemblages, they are unable to compose a truly flexible set of representations of the world or to provide a large functional repertoire to cope with the diversity of a natural environment beyond the niche to which they have evolved. Furthermore, the notion of goal is not explicitly represented. The evolutionary pressure of flexibility leads to the development of other systems (not necessarily one fully general such system) where explicit representation is employed to support goal-oriented reflective behavior which compensates for the shortcomings of the sensorimotor system. Furthermore, it is the mechanism that facilitates the intimate interaction between these two sets of systems that allows humans to achieve a high degree of flexibility and efficiency at the same time.

The same problem also arises in AI, though along quite a different (almost reversed) evolutionary course. In "classical" AI, explicit planning precedes execution, taking the form of a centralized sequential deliberation based on goals

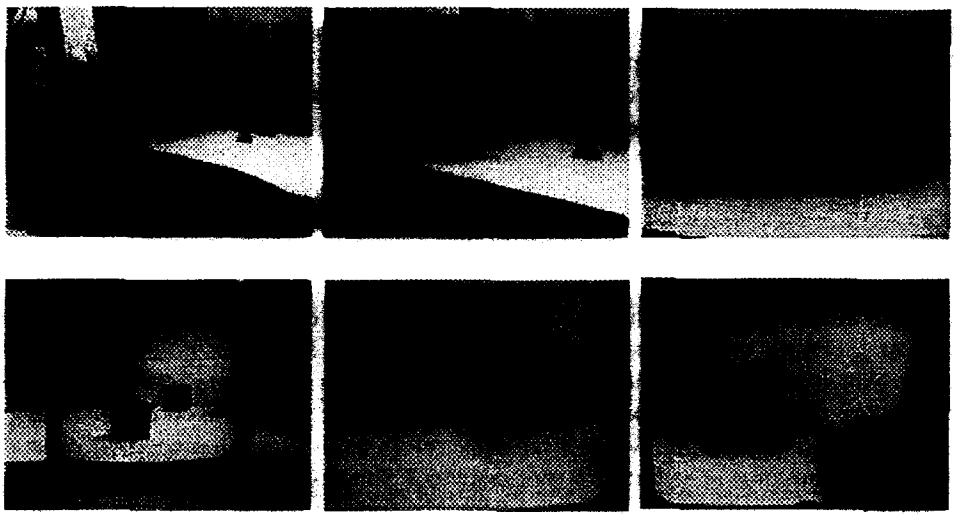

Fig. 7. Estimating depth based on motion parallax. The top shows three snap shots of the robot arm moving around three obstacles (from left to right). The frames in the bottom row show images seen through the robot's eye (the video camera). In this experiment, a large object is placed behind two small objects. The looming neural network is able to detect the small ones first based on motion parallax and guides the robot arm through them first.

and a world model to yield a sequence of actions. Such an approach may impose unrealistic requirements for modeling and perception since all relevant information must be available before planning begins. This makes it hard to adapt the plan to events not predicted by the model. In reaction to this, some critics have advocated *reactive systems* in which selection and execution of actions are inextricably intertwined (e.g., the subsumption architecture of Brooks [9], Kaelbling [24], Agre and Chapman [1]). Such reactive systems are parallel and distributed, with a hardwired priority scheme which is fixed at "compile time". There is no deliberation, and no model. Inhibition/suppression rules determine which models will control action, on the basis of current input. However, such reactive systems are hardwired, difficult to design, and completely data-oriented, and the goals are implicit and thus cannot be referred to in decision making. This leads to the development of hybrid systems to achieve a dynamic intertwining of planning and execution [35–37, 42] (see also [46] on dynamic planning). Here, though along an almost reversed evolutionary course, we see the need for reactive (instinctive) behavior and reflective behavior and, moreover, their integration into a coherent whole.

In the next section, we will outline a computer vision system constructed within the framework of schema theory to demonstrate how the integration of planning and reaction can be achieved in a manner consistent with the principles of cooperative computation and evolution by modulation. Before we do that, however, an important contrast between biological evolution and that of AI is

worth noting. With its roots in symbolic representation, the AI tradition is carried over from planning to reactive systems and into the latest development of the hybrid approach of dynamic planning. In biological evolution, on the other hand, we see two distinct representational schemes subserving instinctive and reflective behaviors, namely, implicit (subsymbolic) and explicit (both symbolic and subsymbolic) representations. Using implicit representations, the animal achieves efficiency in "routine" performance of instinctive behaviors like feeding or fleeing, whereas explicit representation lends itself to contemplation and, through dynamic composition, provides a rich set of representations of the world as the basis for flexible reasoning and planning. The process of learning to drive a car with the help of verbal instruction shows that explicit representations may orchestrate a variety of implicit representations (such as visual recognition of an impending collision) and that constant rehearsal of explicit representations may yield to their automatization in which they come to form schemas that are normally under implicit control, while still possessing paths for explicit access in exceptional circumstances.

The distinction between the reflective and instinctive behaviors can be illustrated in the cortical visual systems of primates, i.e., the "what" pathway which performs object recognition and the "how" pathway which mediates interaction with objects [18, 45]. The "what" pathway runs along the ventral part of the brain whereas the "how" system involves the dorsal part (with massive interconnection between them along the way). Lesion to the "what" pathway in monkey impairs the performance of visual pattern discrimination and recognition but not object location tasks. Quite the opposite results were observed in monkeys with lesions to the "how" pathway. Further observations have been reported in human patients. Goodale and Milner [18] have noted the ability of a patient with a ventral lesion to carry out a variety of object manipulations even though unable to demonstrate explicit knowledge of the object parameters used to guide these actions. For example, when asked to pick up objects with various sizes or orientations, the patient is able to preshape the hand according to the size and orientation of the object to pick it up. Yet, when asked to indicate the size or the orientation of the objects either by means of her hand or verbally, the patient failed to do so. The opposite deficit was observed in another patient with lesion to the "how" pathway. While this patient had no difficulty in recognizing objects, her ability to pick up such objects is severely impaired. Further review of such data, and a fuller expansion of the argument in Section 7, may be found in [5].

The advantage of using the subsymbolic representation is efficiency. We have seen examples of such representation in previous sections, e.g., the retinotopic maps in the frog's for coding the location and direction of a looming stimulus or the potential field in detour behavior. The utilization of subsymbolic representations is not exclusive to instinctive behavior; indeed, as we saw in our discussion of learning to drive, instances of such representation can be found even in behaviors that are considered highly cognitive. Consider the process of acquiring a second language. One first learns a set of grammatical rules and uses the language poorly by applying those rules. Speaking a new language (or even one's

native language, of that matter) in such a way is inefficient. However, as learning progresses and one's mastery of the language improves, more and more of the rules "disappear" from consciousness and one starts to use the language without explicitly thinking about it. Only after such transferring of the rules (or knowledge in more general situations) from explicit to implicit representation, so that the agent becomes capable of behaving without explicitly reasoning on the rules (or symbolic knowledge), can efficiency be achieved.

We have charted the nearly opposite evolutionary courses in biology and AI and shown that, in the end, the agents (both biological and artificial) come to be confronted with the same set of problems due to the demands of efficiency and flexibility in interaction with the environment which lead to the development of reflective and instinctive behaviors and the corresponding underlying representational schemes. We stress that it is the integration of these two behaviors that brings out the essence of intelligence.

7. From vision to distributed planning

Our study of animal-based robot design introduces an evolutionary basis for the design of intelligent systems. It is our task now to chart the "evolutionary breakthrough" which adds "reflective" capabilities to this basis in a fashion consistent with the principles of cooperative computation and evolution by modulation that we have enunciated above—but, to conserve space, we will do this very briefly, referring the reader to Arbib and Goodale [5] for details. We do this by addressing the integrative question of how complex visual scenes affect human behavior. The approach to schema-based interpretation in the VISIONS computer vision system [14] employs active, independent, schema instances, and the schemas encode mechanisms for using features in multiple representations, conducting information fusion and sensor fusion in a knowledge-directed manner. The knowledge required for interpretation is stored in LTM (long-term memory) as a network of perceptual schemas, while the state of interpretation of the particular scene unfolds in STM (working memory) in the form of a network of schema instances. Each schema instance has an associated activity level (or confidence level) which changes on the basis of interactions with other units in the (dynamically reconfigurable) STM network. The STM network makes context explicit: each object represents a context for further processing, using to advantage the relations among objects. When a schema instance is activated, it is with an associated area of the image and an associated set of local variables. Different instances of a given schema may be associated with separate portions of the image if they correspond to different instances of the "object" the schema represents. The structure of STM is further constrained in part by relationships encoded within LTM, both those between schemas for inter-object relations and those within a schema for geometric relations of parts.

In the VISIONS system, it is the user who starts the interpretation process by

invoking general goals such as "interpret this image as a road scene" or such specific goals as "find the sidewalk in this image". The point we stress here is that the above considerations still have much to offer when applied to analyzing the vision of a robot or an animal. However, in such systems, it is the current goals of the autonomous system, not the demands of some user, that guide the process of *action-oriented* perception. Changing goals change the perceptual demands as action proceeds.

The VISIONS system exemplifies a distributed form of planning—as the result of activity emerging in a flexible network—involving generativity to form novel patterns of schema activation that may involve creation of novel "networks". However, if we extend the analysis so that motor schemas are activated as well as the perceptual schemas of STM, we see that planning will be intertwined with execution, with patterns of schema activation modified and thus the "plan" updated as action proceeds and new sensory stimulation is obtained (similar principles are adapted in the IPUS architecture for acoustic signal interpretation [28], see also [41] for active vision). We thus see planning as a process emerging from the cooperative computation of multiple agents rather than being imposed by a separate executive planning system, thus providing an approach to the reactive (we prefer the term "dynamic") planning reviewed by Lyons and Hendriks [35].

Schema theory provides a distributed model of computation, supporting many concurrent activities for recognition of objects, and the planning and control of different activities. Each instance of a schema has an associated *activity level*. That of a perceptual schema represents a "confidence level" that the object represented by the schema is indeed present; while that of a motor schema may signal its "degree of readiness" to control some course of action. The use, representation, and recall of knowledge is mediated through the activity of a network of interacting computing agents, the schema instances, which between them provide processes for going from a particular situation and a particular structure of goals and tasks to a suitable course of action (which may be overt or covert, as when learning occurs without action or the animal changes its state of readiness). This activity may involve passing of messages, changes of state (including activity level), instantiation to add new schema instances to the network, and deinstantiation to remove instances. Moreover, such activity may involve self-modification and self-organization, but the topic of learning (save by implication in our analysis of evolutionary design) is beyond the scope of this article (readers are referred to the extension by Earl and Firby to construct new schemas for reactive planning). A schema network does not, in general, need a top-level executor since schema instances can combine their effects by distributed processes of competition and cooperation (i.e., interactions which, respectively, decrease and increase the activity levels of these instances), rather than the operation of an inference engine on a passive store of knowledge. This may lead to apparently emergent behavior, due to the absence of global control.

The transition we have seen from frog visuometer coordination to human visual

perception[4] is reminiscent of the "Great Move" charted by Newell [39] from using specialized materials to support different schemas to using a medium in which it is possible to compose copies of whatever schemas are needed to form novel representations. Contrasting the study of ethologists of the adaptive character of lower organisms with the rich repertoire of human cognitive abilities, Newell observes that:

> Finding feasible representations gets increasingly difficult with a richer and richer variety of things to be represented and richer and richer kinds of operational transformations that they undergo. More and more interlocking representation laws need to be satisfied. . . . Instead of moving towards more and more specialized materials with specialized dynamics to support an increasingly great variety and intricacy of representational demands, an entirely different turn is possible. This is the move [the Great Move] to using a neutral, stable medium that is capable of registering variety and then *composing* whatever transformations are needed to satisfy the requisite representation law. Far from representational constriction, this path opens up the whole world of indefinitely rich representations. (Newell [39, p. 61]).

We may note here that Brooks [9] sets an ethologically inspired hierarchy of levels of control (mentioned earlier as being in the spirit of our schema-based design principles), each biasing rather than replacing the one below it, in opposition to the "classical" view of abstract operators applied to uniform representations. In his general cognitive architecture for problem solving, SOAR, Newell offers a somewhat broader version of classical AI, since he allows a variety of problem spaces—but nonetheless sees these each as being implemented in some uniform medium. However, it is mistaken to see this as a sharp dichotomy in which one school or the other must prove triumphant. The schema theoriest (as in our discussion of the "what" and "how" visual systems) explains a complex cognitive function through the interaction of "instinctive" schemas, implemented in specifically evolved circuitry, and "abstract" schemas that are developed through learning and experience in relatively "general purpose" (highly adaptive, post-Great-Move) circuitry. An intelligent system needs to combine the ability to react rapidly (jumping out of the way of an unexpected vehicle when crossing the street) with the ability to abstractly weigh alternatives (deciding on the best way to get to the next appointment).

---

[4] The VISIONS model is not a biological model and, indeed, the *generic* architecture sketched above allows an unlimited number of schema instances to process simultaneously—ignoring crucial issues of focus of attention. However, Arbib and Goodale [5] review evidence on primate cerebral cortex to suggest how processing algorithms and memory strategies inspired by the distributed planning model outlined so briefly above may indeed be encoded in the brain. The result is not a brain model *per se*, but rather a sketch of a dramatically new approach to modeling the interaction of cortical mechanisms for vision and action.

In summary, a satisfactory account of Newell's "Great Move" should not seek a complete break from using specialized materials to support different schemas to using a medium in which it is possible to compose copies of whatever schemas are needed to form novel representations. Rather, we should provide—in the manner of schema theory—insight into how instinctive behavior provides a basis for, and is intertwined with, rational behavior. When we study frogs, we see the animal's behavior mediated by the dynamic interaction of multiple special-purpose schemas implemented in dedicated neural circuitry. But when we seek to understand human vision, we combine a model of low-level vision implemented across a set of dedicated brain regions [12] with a variety of "general-purpose" media in which copies of schemas (schema instances) can be assembled, parameterized, and bound to regions of the image as they compete and cooperate in a process of distributed planning which creates an interpretation of a visual scene (cf. [3], especially Sections 5.1, 5.2, 7.3, and 7.4). The contrast between frog visuomotor coordination and the flexibility of human visual perception makes explicit the contrast between those schema assemblages that are "evolutionarily hardwired" into patterns of competition and cooperation between specific brain regions, and those which can, through multiple instantiations (both data- and hypothesis-driven), yield totally novel forms to develop new skills and represent novel situations.

Principle 4 (*Interpretation as planning*). The mechanisms of distributed interaction seen in visual interpretation extend to distributed goal-directed planning.

Principle 5 (*Distributed goal-directed planning*). Moreover, the "Great Move" takes us from dedicated circuitry for every schema to the creation of novel schemas distributed across relatively general-purpose machinery. Distributed goal-directed planning combines reactive aspects (dynamic updating) with working memory and focus of attention mechanisms.

The hedge "relatively" in the expression "relatively general-purpose machinery" is important. Recent advances in neuroanatomy have increasingly demonstrated that what were once thought to be large undifferentiated brain regions can in fact be discriminated into far smaller regions with distinctive patterns of input and output connections [19]. Presumably, each of these regions can access a distinct set of other regions, and as these sets of subsets can be developed hierarchically, the result can be a "general-purpose" representation—but one which is distributed across more or less specialized partial representations (in the spirit of Principle 3) each in a specified brain region, rather than being a set of states of neutrally-addressed registers in the totally uniform computational medium of the serial general-purpose computers used in Newell's many contributions to "classical" AI.

8. Relating structure and function: from biology to technology

Many authors have emphasized that the brain can be analyzed at many different levels of detail going all the way from the overall brain through the various anatomically separable brain regions to layers or modules of cells down to the individual neurons, and from there down even further to various cellular components and even to the very molecules themselves. This corresponds to the right-hand path shown in Fig. 8(a). However, what this figure emphasizes is that in computational neuroscience we will as often start from a behavior of the organism as from a concern with particular brain regions of the organism. We must stress that a functional unit (a schema) must not be equated with a structural unit (a component or processor)—in general, a schema may be implemented across several components, and a component may contribute to several functions (Fig. 1(b)). When we start from the behavior we are committed to a functional analysis; when we start from one or more brain regions, we are committed to a structural analysis. In this paper we have given some sense of how schemas may be defined without any commitment as to their implementation within specific neural circuits, but we have then suggested how this functional analysis may be considered by the data of neurobiology. We can directly confront the functional decomposition of schemas with the structural decomposition of brain regions, etc., by using lesion analysis to see whether our account of interacting schemas when coupled with hypotheses about which particular brain regions are involved (recall Fig. 1) bear up when we look at the behavior of animals with brain lesions. However, structure and function may be brought even more directly together, as we saw in Section 4, when we analyze the neural circuitry in a brain region to see

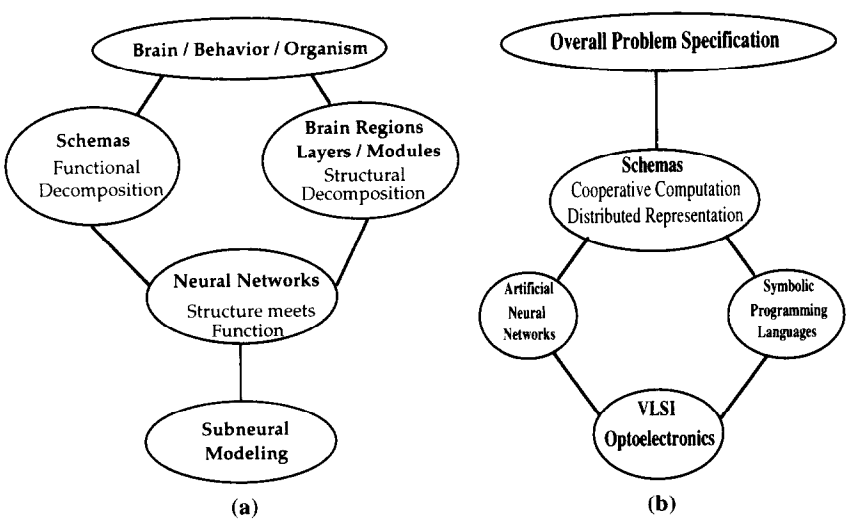

Fig. 8. Views of level of analysis of (a) brain and behavior and (b) a distributed AI system, highlighting the role of schemas as an intermediate level of functional analysis in each case.

whether it can provide the functionality demanded by the schemas mapped on to this region. Our examples from frog visuomotor coordination have shown both that a schema inferred on the basis of a functional analysis may indeed correspond to the interaction of many brain regions, and that a specific brain region may be implicated in a number of different schemas. Finally, although we have not discussed this above, Fig. 8(a) emphasizes that neuroscience often proceeds to the subneural level to explain how, for example, the properties of synapses can mediate the computations and learning properties of individual cells (see, e.g., Section 8.1 of [3]).

When we turn from the brain to artificial systems—whether under the rubric of distributed artificial intelligence, neural engineering, or perceptual robotics—we find a somewhat different situation, as shown in Fig. 8(b). In the study of the brain we are committed to the fact that any functionality must—at least in principle (though we may choose to stop our modeling of at least some functions at the schema level)—be played over neural networks. In contrast, in an artificial system there is no such commitment. We pick the optimal implementation for a particular functionality and this may indeed differ from function to function. Thus, rather than seeing schemas as providing a path alternate to that of the structural path, as we did for computational neuroscience in Fig. 8(a), we here see schemas as providing the language of cooperative computation and distributed representation in which we provide the coarse-grain analysis of the interacting functionalities required to meet the overall problem specification. Branching occurs from the schema level to the level of fine-grain implementations—Fig. 8(b) shows just two examples, namely artificial neural networks and symbolic programming. The unification may then occur at the lowest level where, irrespective of the intermediate level of implementation, we come to a computational medium such as VLSI or optoelectronics which may either be used uniformly as a substrate for many different intermediate forms, or may be chosen varyingly in the design of dedicated processors for different families of schemas. In the spirit of Principle 5, it will be a major design issue to determine which schemas are most efficiently implemented in special-purpose hardware, and which—and in which various subsets—are programmed to share general-purpose hardware with a variety of other schemas.

References

[1] P.E. Agre and D. Chapman, Pengi: an implementation of a theory of action, in: *Proceedings AAAI-87*, Seattle, WA (1987) 268–272.

[2] M.A. Arbib, Levels of Modeling of Visually Guided Behavior (with peer commentary and author's response), *Behav. Brain Sci.* **10** (1987) 407–465.

[3] M.A. Arbib, *The Metaphorical Brain 2: Neural Networks and Beyond* (Wiley-Interscience, New York, 1989).

[4] M.A. Arbib, Schema theory, in: S. Shapiro, ed., *The Encyclopedia of Artificial Intelligence* (Wiley-Interscience, New York, 2nd ed., 1992) 1427–1443.

[5] M.A. Arbib and M.A. Goodale, Visual routes to action and knowledge: a new perspective on the cognitive architecture of vision (to appear).

[6] M.A. Arbib and D.H. House, Depth and detours: an essay on visually-guided behavior, in: M.A. Arbib and A.R. Hanson, eds., *Vision, Brain, and Cooperative Computation* (Bradford Books/MIT Press, Cambridge, MA, 1987) 129–163.
[7] R.C. Arkin, Neuroscience in motion: the application of schema theory to mobile robotics, in: J.-P. Ewert and M.A. Arbib, eds., *Visuomotor Coordination: Amphibians, Comparisons, Models, and Robots* (Plenum, New York, 1989) 649–671.
[8] V. Braitenberg, *Vehicles: Experiments in Synthetic Psychology* (Bradford Books/MIT Press, Cambridge, MA, 1984).
[9] R.A. Brooks, A robust layered control system for a mobile robot, *IEEE Trans. Rob. Automation* **2** (1986) 14–23.
[10] F. Cervantes-Pérez, R. Lara, and M.A. Arbib, A neural model of interactions subserving prey-predator discrimination and size preference in anuran amphibia, *J. Theor. Biology* **113** (1985) 117–152.
[11] A. Cobas and M.A. Arbib, Prey-catching and predator-avoidance in frog and toad: defining the schemas, *J. Theor. Biol.* **157** (1992) 271–304.
[12] E.A. DeYoe and D.C. Van Essen, Concurrent processing streams in monkey visual cortex, *Trends Neurosci.* **11** (5) (1988) 219–226.
[13] R.L. Didday, A model of visuomotor mechanisms in the frog optic tectum, *Math. Biosci.* **30** (1976) 169–180.
[14] B.A. Draper, R.T. Collins, J. Brolio, A.R. Hanson, and E.M. Riseman, The schema system, *Int. J. Comput. Vision* **2** (1989) 209–250.
[15] J.-P. Ewert, Neuroethology of releasing mechanisms: prey-catching in toads, *Behav. Brain Sci.* **10** (1987) 337–405.
[16] A.H. Fagg, M.A. Lewis, T. Iberall, and G. Bekey, R^2AD: Rapid Robotics Application Development environment, in: *Proceedings IEEE Conference of Robotics and Automation* (1991) 1420–1426.
[17] J.J. Gibson, The optical expansion-pattern in aerial location, *Am. J. Psychol.* **68** (1955) 480–484.
[18] M.A. Goodale and A.D. Milner, Separate visual pathways for perception and action, *Trends Neurosci.* **15** (1992) 20–25.
[19] P. Goldman-Rakic, Parallel systems in the cerebral cortex: the topography of cognition, in: M.A. Arbib and J.A. Robinson, eds., *Natural and Artificial Computation* (MIT Press, Cambridge, MA, 1991) 155–176.
[20] I. Horswill, Analysis of adaptation and environment, *Artif. Intell.* **73** (1995).
[21] D.H. Hubel and T.N. Wiesel, Receptive fields, binocular and functional architecture in the cat's visual cortex, *J. Physiol. (London)* **160** (1962) 106–154.
[22] D. Ingle, Spatial vision in anurans, in: K.V. Fite, ed., *The Amphibian Visual System* (Academic Press, New York, 1976) 119–140.
[23] D. Ingle and K. van Hoff, Visually elicited evasive behavior in frogs: giving memory research an ethological context, *BioSci.* **40** (4) (1990) 284–291.
[24] L.P. Kaelbling, An architecture for intelligent reactive systems, in: *Proceedings Workshop on Planning and Reasoning about Action*, Timberline, OR (1986) 235–250.
[25] O. Khatib, Real-time obstacle avoidance for manipulators and mobile robots, *Int. J. Rob. Res.* **5** (1986) 90–98.
[26] J.-C. Latombe, *Robot Motion Planning* (Kluwer Academic Publishers, Dordrecht, Netherlands, 1991).
[27] D.N. Lee and J.R. Lishman, Visual control of locomotion, *Scand. J. Psychol.* **18** (1977) 224–230.
[28] V.R. Lesser, S.H. Nawab and F.I. Klassner, IPUS: an architecture for the integrated processing and understanding of signals, *Artif. Intell.* **76** (1995).
[29] J.Y. Lettvin, H. Maturana, W.S. McCulloch, and W.H. Pitts, What the frog's eye tells the frog brain, *Proc. IRE* **47** (1959) 1940–1951.
[30] J.-S. Liaw, Visuomotor coordination in anurans, mammals, and robots, Ph.D. Thesis, Department of Computer Science, University of Southern California, Los Angeles, CA (1993).
[31] J.-S. Liaw and M.A. Arbib, Neural mechanisms underlying direction-selective avoidance behavior, *Adaptive Behav.* **1** (1993) 227–261.

[32] J.-S. Liaw, A.H. Fagg, and M.A. Arbib, Robot obstacle avoidance using a biologically-based looming perception neural network (to appear).
[33] J.-S. Liaw, I.K. King, and M.A. Arbib, Visual perception of translational and rotational motion in neural networks, in: M. Omidvar and R. Mohan, eds., *Vision, Progress in Neural Networks* (Ablex, Norwood NJ, to appear).
[34] D.M. Lyons and M.A. Arbib, A formal model of computation for sensory-based robotics, *IEEE Trans. Rob. Automation* **5** (1989) 280–293.
[35] D.M. Lyons and A.J. Hendriks, Planning, reactive, in: S. Shapiro, ed., *The Encyclopedia of Artificial Intelligence* (Wiley-Interscience, New York, 1992) 1171–1181.
[36] D.M. Lyons and A.J. Hendriks, Exploiting patterns of interaction to achieve reactive behavior, *Artif. Intell.* **73** (1995).
[37] P. Maes, A bottom-up mechanism for behavior selection in an artificial creature, in: S. Wilson and J.-A. Meyer, eds., *Simulation of Animal Behavior: From Animals to Animats* (MIT Press, Cambridge, MA, 1991) 238–246.
[38] B.W. Mel, *Connectionist Robot Motion Planning: A Neurally-Inspired Approach to Visually-Guided Reaching* (Academic Press, Boston, MA, 1990).
[39] A. Newell, *Unified Theories of Cognition* (Harvard University Press, Cambridge, MA, 1990).
[40] W.H. Pitts and W.S. McCulloch, How we know universals, the perception of auditory and visual forms, *Bull. Math. Biophys.* **9** (1947) 127–147.
[41] Y. Roth and R. Jain, Knowledge caching for sensor-based systems, *Artif. Intell.* **71** (1994) 257–280.
[42] M. Schoppers, The use of dynamics in an intelligent controller for a space faring rescue robot, *Artif. Intell.* **73** (1995).
[43] O.G. Selfridge, Pandemonium: a paradigm for learning, in: *Mechanisation of Thought Processes* (Her Majesty's Stationery Office, London, 1959) 511–531.
[44] J.L. Teeters and M.A. Arbib, A model of anuran retina relating interneurons to ganglion cell responses, *Biol. Cybern.* **64** (1991) 197–201.
[45] L.G. Ungerleider and M. Mishkin, Two cortical visual systems, in: D.J. Ingle, M.A. Goodale and J.W. Mansfield, eds., *Analysis of Visual Behavior* (MIT Press, Cambridge, MA, 1982).
[46] R. Vijaykumar, Y. Liu, and M.A. Arbib, Dynamic planning for sensor-based robots, in: *Proceedings, SYROCO (Symposium on Robot Control)* Barcelona, Spain (1985) 401–406.
[47] W.G. Walter, *The Living Brain* (Duckworth, London, 1953).
[48] D.L. Wang and M.A. Arbib, Modeling the dishabituation hierarchy: the role of the primordial hippocampus, *Biol. Cybern.* **67** (1992) 535–544.
[49] A. Weitzenfeld, NSL, Neural Simulation Language, Technical Report 91-05, Center for Neural Engineering, University of Southern California, Los Angeles, CA (1991).